

Overview: Open data, transparency and campaigning

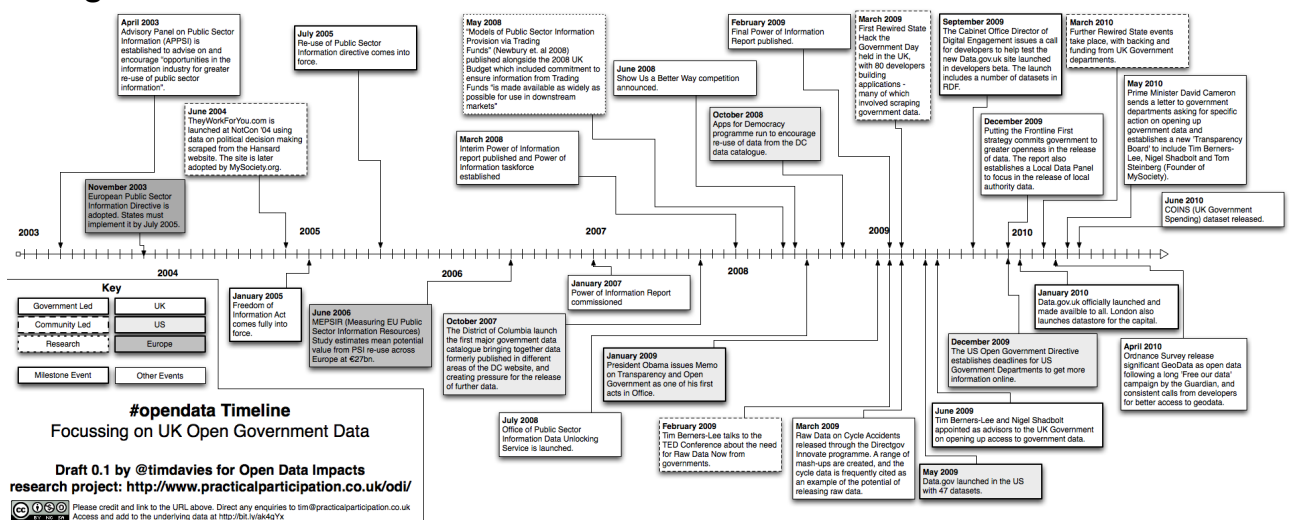
Brief notes by Tim Davies, Director, Practical Participation

tim@practicalparticipation.co.uk / @timdavies / <http://www.timdavies.org.uk>. Questions, comments and ideas around possible training and development workshops / ways of exploring further welcome.

Quick links

Search for info on your MP	Find government data	Request more data
http://www.theyworkforyou.com	http://data.gov.uk	http://www.opsi.gov.uk/unlocking-service/opsipage.aspx
http://www.publicwhip.org	http://guardian.org.uk/data-store	http://www.whatdotheyknow.com/
http://election.theyworkforyou.com	http://ckan.net	http://openup.tso.co.uk

Background & Context



The last 18-months have seen an impressive array of policy initiatives and practical actions leading to the release of datasets from governments in the UK, the US and across the world in open and re-usable formats online. Datasets ranging from the location of educational institutions, to details of taxation and government spending, have been brought together in data portals such as data.gov and data.gov.uk.

The open government data 'movement' has three broad constituent parts:

- **An open Public Sector Information (PSI) movement** - drawing upon economic arguments to call for government data to be released and made freely re-usable. Often drawing upon comparisons between EU context where government collected data is copyright and restricted, and the US where government datasets are more open and large industries have developed on the back of them (e.g. Weather data; Geodata etc.)
- **A transparency movement** - linked to Access to Information and Freedom of Information movements - calling for the release of data in the interests of democratic

empowerment.

- **Digital government & semantic web computerization movements** - focussed on the potential for innovation and more efficient working when data is made available for computer processing: and to build open networks of knowledge across the Internet.

Many different groups can be found within the open government data 'movement' - from groups calling for aid transparency, to SME companies seeking to address what are seen as unfair data monopolies.

Policy context:

- The <http://data.gov> initiative in the US proceeded from Obama's first executive order on taking power as President.
- <http://data.gov.uk> in the UK was initiated by Gordon Brown in 2009.
- Since coming to power in 2010 the Coalition Government in the UK have continued to push open data initiatives - thought with a slightly different 'transparency' and 'accountability' framing.
 - A requirement has been placed on local authorities to publish all spending over £500 by January, listing supplier and spend.
 - Government departments are under a similar requirement for all spend over £25,000, and have been asked to publish senior staff pay details and internal organizational diagrams.
 - Francis Maude has spoken of the need for a 'Freedom of Data' act, and has called for all responses to Freedom of Information requests that contain data to provide that data in machine-readable forms (i.e. Excel spreadsheet rather than print-out of PDF files...)
 - Aid Transparency has been high up the government's development agenda.
- The World Bank have released significant amounts of their data as open data.
- Australia, New Zealand and many European countries have ongoing open data initiatives and campaigns.

Beyond government data

It's not only government supplied data that is of interest to campaigners:

- Projects like TheyWorkForYou.com and PublicWhip.org generate structured data about politicians voting records by 'scraping' parliamentary records;
- Data Journalists (led by innovators at The Guardian amongst other places) publish their research as open accessible spreadsheets of data that others can re-use.
- Some NGOs and community organisations are publishing open datasets.

Why data?

One of the key properties of data is that it can be easily manipulated by computer - allowing datasets to be combined, visualized, explored and used in many more ways than a written report or printed document can.

Where to find data

For official government data - the guardian's World Government Data Search looks across a range of data catalogues like <http://data.gov.uk>. Find it at <http://www.guardian.co.uk/world-government-data> and search for keywords or topics of interest to you.

You can also search <http://data.gov.uk> direct to browse data by department or topic.

<http://ckan.net/> provides a catalogue of open data from many different sources - including government data, NGOs and research projects. It is a good place to 'register' any open data you create. It is also wiki-like, meaning any user can edit the records - allowing the creation of 'collections' of data on a particular topic: e.g. 'arms trade'.

ScraperWiki.com provides a collection of 'scrapers' which collect structured data from unstructured data-sources (i.e. make open data where the original publisher didn't provide it). For example, generating **a dataset of hospitality** received by UK Government Ministers, originally only available as a large collection of different word documents is now here: <http://scraperwiki.com/scrapers/government-meetings-with-external-organisations-ne/> and available for download.

If you are looking for a particular dataset - it can be worth asking in the data.gov.uk forums, or using the #opendata hash-tag on Twitter.

Data on MPs and voting records is available from www.theyworkforyou.com in the UK, and the PublicWhip.org project collects more detailed voting records and makes them available.

When data isn't available

Try using the Public Data Unlocking Service to request that data is proactively published: <http://www.opsi.gov.uk/unlocking-service/opsipage.aspx?page=unlockindex>

If using the Freedom of Information Act to request data, remind the recipient of Francis Maude's policy statements on the need to provide machine-readable data in return: http://www.conservatives.com/News/Speeches/2010/10/Francis_Maude_3bn_of_efficiency_savings.aspx

If the information is available on websites, but not as structured data - consider putting a request on <http://www.scraperwiki.com> for someone to build a tool to screen-scrape the data.

Consider using any of the 'data competitions' (e.g. <http://openup.tso.co.uk>) as a higher-profile way to ask for a dataset: emphasizing the government's focus on accountability through transparency in other sectors such as local authority spending and aid.

Use the facts you can find from datasets like COINS (<http://data.gov.uk/dataset/coins>) to better structure Freedom of Information requests or crowdsourcing activities.

Explore ways to 'crowd-source' the data by calling on campaigners and supporters to find out particular facts - and to enter them into shared online spreadsheets (e.g. using Google Spreadsheets and Google Forms you can create an easy way for people to collaboratively input into a shared document - which can be instantly published online). Crowdsourcing tools like Ushahidi can also be used to develop projects such as <http://WhereAreTheCuts.org> - crowdsourcing reports of public spending cuts.

Working with data

Working with data scares many people - but it can start off very simply:

1. **Using data-driven websites** such as <http://TheyWorkForYou.com> (MPs speeches and voting) or <http://WhereDoesMyMoneyGo.com> (government spending) which have taken government data and made it available in more accessible forms.
2. **Downloading and exploring a single dataset** - many datasets can be opened in spreadsheet software like Excel. Sort and filter the columns to look for interesting information.
3. **Visualise the data** - using a tool like IBM Many Eyes where you can upload simple datasets and explore a range of different ways of presenting the data.
4. **Building a mash-up** - using tools like Google Spreadsheets and Google Fusion Tables (search for an use online), or Freebase Gridworks (available for free download) to explore and combine datasets.

Google Fusion Tables will allow you to upload any spreadsheet, and, if it contains place names, quickly 'geocode' the data for displaying on a map. You can also combine two datasets - matching on any shared keys (e.g. MP name; Town name; Constituency) to build larger datasets.

5. **Holding a hack day** - hack days like those organized by <http://www.rewiredstate.org> bring together developers (coders/geeks) and people with problems to solve and spend one or two days of concerted effort creating 'hacks' (rapid prototypes) which address those issues, often using open data.

For example, a hack-day could look to generate visualizations concerning arms licenses, or to create tools that support campaigners to get information to use when writing to MPs.

6. **Commissioning open data-based tools** - developing hack-day created prototypes, or other ideas, into full working tools.
7. **Training activists in using data** - through workshops and hands-on activities.
8. **Releasing datasets** - from in-house research or crowd-sourced data - and inviting supporters to use the data in creative ways. For example, putting researched data into Google Spreadsheets and, much as the Guardian Datablog does (<http://www.guardian.co.uk/news/datablog>), sharing links to that data whenever posting news stories or website pages based upon it.

Going further

Search for the #opendata community on Twitter; or the 'Open Government Data' mailing lists run by Open Knowledge Foundation. Most of the links above will also provide access to further practical and background information on open government data.

Tim Davies, Practical Participation (tim@practicalparticipation.co.uk) can offer consultancy, training, workshops and support for organisations exploring the use of open data in campaigning. Please do get in touch to explore more...